

# THE EFFECTS OF A MODEL STATEMENT

## The Effects of a Model Statement on Information Elicitation and Deception Detection in Multiple Interviews

Haneen Deeb, Aldert Vrij, and Sharon Leal

University of Portsmouth, Department of Psychology, United Kingdom

**ACCEPTED IN ACTA PSYCHOLOGICA, APRIL 2020**

### Author Note

Correspondence concerning this article should be addressed to Haneen Deeb, Department of Psychology, University of Portsmouth, King Henry I Street, PO1 2DY, Portsmouth, United Kingdom. Telephone: 0044 (0) 23 9284 6318. Fax: 0044 (0) 23 9284 6300. Email: [haneen.deeb@port.ac.uk](mailto:haneen.deeb@port.ac.uk)

**Haneen Deeb:** Conceptualisation, Software, Investigation, Validation, Formal analysis, Data curation, Writing – Original Draft, Visualisation, Project administration

**Aldert Vrij:** Conceptualisation, Methodology, Formal analysis, Writing – Review & Editing, Supervision, Funding acquisition

**Sharon Leal:** Conceptualisation, Methodology, Writing – Review & Editing

## Abstract

Lie detection researchers started developing interview techniques to enhance deception detection in forensic settings. One of those techniques is the Model Statement, which has been shown to be effective for eliciting information and cues to deception in single interviews. In the current research, we focused on the effect of the Model Statement in multiple interviews. Participants ( $N = 243$ ) were interviewed three times—each time one week apart—about a genuine (truth tellers) or fabricated (lie tellers) memorable event. They listened to a Model Statement at Time 1, Time 2, Times 1 and 2, or not at all. Hypotheses focused on participants' verbal reports at Time 3 and on unique details provided across the three interviews. In both instances, truth tellers provided more core and total details and complications and fewer common knowledge details and self-handicapping strategies and obtained higher proportion scores of (i) complications and (ii) core details than lie tellers. Complications and proportion of complications were the most diagnostic cues. The Model Statement was effective only when presented at Time 1, resulting in more common knowledge details. No Veracity x Model Statement interaction effects emerged.

*Keywords:* deception; Model Statement; multiple interviews; core details; peripheral details; complications

### Highlights

- When using the Model Statement interview technique in three interviews, truth tellers provided more core and total details and complications and fewer common knowledge details and self-handicapping strategies and obtained higher proportion scores of (i) complications and (ii) core details than lie tellers.
- Complications and proportion of complications were the most diagnostic cues.
- The Model Statement was effective only when presented at Time 1, resulting in more common knowledge details.

The Effects of a Model Statement on Information Elicitation and Deception Detection  
in Multiple Interviews

## 1. Introduction

In recent years, interviewing techniques were developed in which the interviewer plays an active role in enhancing verbal differences between truth tellers and lie tellers. These techniques typically focused on single interviews (e.g., Leal, Vrij, Warmelink, Vernham, & Fisher, 2015; also see Vrij et al., 2017). However, suspects in forensic and intelligence contexts are often interviewed multiple times, yet research on how multiple interviews affect information elicitation and deception detection is scarce (Granhag & Strömwall, 2002; Harvey, Vrij, Leal, Lafferty, & Nahari, 2017). In this paper, we examine the effects of one technique, playing an audiotope of a Model Statement at the beginning of the interview, on information elicitation and deception detection in multiple interviews. The study was pre-registered on [anonymous link:

[https://osf.io/k6az7/?view\\_only=e7a6146709b342e28be7353fd3252abe](https://osf.io/k6az7/?view_only=e7a6146709b342e28be7353fd3252abe)].

### *1.1. The Model Statement interview technique*

Research has shown that cues to deception are typically faint and unreliable (DePaulo et al., 2003). However, the use of specific interview approaches, such as the Cognitive Credibility Assessment approach that focuses on differences in cognitive processing and strategy use between truth tellers and lie tellers, was shown to be effective for lie detection purposes (Vrij, Fisher, & Blank, 2017; Vrij, Fisher, Blank, Leal, & Mann, 2016; Vrij, Leal, Mann, Vernham, Brankaert, 2015). One component of this approach is the Model Statement interview technique (Leal et al., 2015). A Model Statement is an example of a detailed statement unrelated to the topic of investigation. Before suspects are interviewed, they listen to an audiotope of the Model Statement and are encouraged to provide an account that is as

rich in detail as the statement they heard. The Model Statement raises interviewees' expectations about the amount of detail they are expected to provide (Ewens et al., 2016; Festinger, 1954). Therefore, interviewees who listen to the Model Statement provide more information than interviewees who are solely instructed to report as many details as possible (Leal et al., 2015). This may be explained by the fact that the Model Statement is an example, which is easier to comprehend than the general instruction to report many details (Vrij, Leal, & Fisher, 2018).

The Model Statement typically results in both truth tellers and lie tellers providing more information in interviews (see Vrij, Leal, & Fisher, 2018 for an overview). However, differences emerge when the type of detail they report is taken into account. After listening to the Model Statement, truth tellers' reports sounded more plausible than lie tellers' reports (Leal et al., 2015) and included more complications (Vrij, Leal, et al., 2017) and verifiable details (Harvey, Vrij, Leal, et al., 2017). In contrast, lie tellers' reports included more peripheral details, common knowledge details and self-handicapping strategies than truth tellers' reports (Leal, Vrij, Deeb, & Jupe, 2018; Vrij, Leal, Jupe, & Harvey, 2018; Vrij, Leal, et al., 2017). In the current study, we attempted to distinguish lie tellers and truth tellers based on the number of core details, peripheral details, common knowledge details, self-handicapping strategies, and complications in their verbal statements.

Core details are details that are central to the event under consideration, and, if altered, they change the event (Heuer & Reisberg, 1990). All other details that do not change the main event are considered peripheral. For example, if an interviewee discusses his experience of a music festival, the activities that occurred during the festival would be considered core details, whereas, for example, the public transportation means he took to reach the location of the festival and going to a friend's house to sleep following the festival are considered peripheral details. Although deception research findings on core and peripheral details are

still inconclusive, there is a tendency for truth tellers to provide more core details and for lie tellers to provide more peripheral details in their statements (Leal, Vrij, Deeb, et al., 2018; Sakrisvold, Granhag, & MacGiolla, 2017). There are at least two explanations for this. First, the core event is the potentially incriminating part of the statement and lie tellers tend to avoid discussing potentially incriminating details (Granhag & Hartwig, 2008). Second, lie tellers prefer to embed their lies in truthful experiences (Leins, Fisher, & Ross, 2013). Thus, in the festival example above, a deceptive interviewee may discuss a location he knows and honestly report going to a friend's house to sleep. However, he may lie about what happened at the location where the festival took place. The interviewee thus lies about the core event, but all the peripheral information is truthful. As people typically report more details about truthful than false events (Amado, Arce, Fariña, & Vilarino, 2016), lie tellers thus may report more peripheral details and fewer core details than truth tellers.

Common knowledge details involve scripted information that is commonly known by people (e.g., "We visited the Louvre museum in Paris where we saw the Mona Lisa"). Whereas truth tellers have a personal experience of a reported event and thus can report a unique experience (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996), lie tellers who do not have such a personal experience draw upon general knowledge to report the event (Sporer, 2016). Even when lie tellers have personal experiences of the event, they may not report them due to their desire to keep their stories simple (Strömwall & Willén, 2011; Vrij, Leal, Jupe, et al., 2018).

Self-handicapping strategies are justifications why someone cannot provide certain types of information (e.g., "I can't remember; it was a while ago when this happened"). In an attempt to keep their stories simple but credible, lie tellers avoid providing information but at the same time try to justify this so that their account is not perceived to be suspicious (Vrij, Leal, Jupe, et al., 2018).

Complications are details that affect the story-teller and make the story more difficult than necessary (e.g., “The air conditioning was not working properly in the hotel”; Vrij, Leal, Fisher, Mann, Jo, et al., 2019). Making up complications requires imagination and lie tellers may not have adequate imagination to make up details that sound plausible (Köhnken, 2004; Vrij, 2008). In addition, lie tellers prefer to keep their stories simple (Hartwig, Granhag, & Strömwall, 2007), and reporting complications makes the statement more complex.

A proportion score of cues of truthfulness, the proportion of complications (complications / [complications + common knowledge details + self-handicapping strategies]), has been used in previous research. It highly distinguishes lie tellers from truth tellers (e.g., Vrij, Leal, Jupe, et al., 2018; Vrij, Leal, Mann, et al., 2018). Such within-subject measures are preferred by practitioners (Vrij, 2016, 2019) and scholars (Nahari et al., 2019; Nahari & Vrij, 2015), because—among other reasons—they control for individual differences (e.g., some people are more talkative than others) and situational differences (e.g., some events are richer in detail than others; Vrij, Leal, & Fisher, 2018).

### *1.2. Multiple interviews*

Research on the Model Statement has focused on interviewing participants only once, immediately after the event. In real life, however, suspects are often interviewed on multiple occasions over a period of time (Granhag & Strömwall, 1999; Vrij et al., 2017), so it is important to understand the implications of a Model Statement on multiple interviews.

Research examining how time affects truth tellers’ and lie tellers’ responses is sparse. In two studies, truth tellers and lie tellers were interviewed either directly after an event or after a three-week delay (Harvey, Vrij, Hope, Leal, & Mann, 2017; Harvey, Vrij, Leal, Hope, & Mann, 2017). Truth tellers reported fewer details when interviewed after a three-week delay than when interviewed immediately after an event, whereas lie tellers reported an equivalent amount of detail at both retrieval intervals. In other words, truth tellers showed a ‘*forgetting*

*curve*' (Ebbinghaus, 1885) as a result of the delay, thus remembering and reporting less information after three weeks. In contrast, lie tellers revealed a *stability bias*, a pattern of failing to correctly understand the degree to which memory can change over time (Kornell & Björk, 2009). The joint effect of the forgetting curve and stability bias is that the cue 'richness of detail' which is found to be a diagnostic cue to deceit (lie tellers provide fewer details than truth tellers) in immediate interviews (Amado et al., 2016) becomes less diagnostic in delayed interviews (see McDougall & Bull, 2015; and Nahari, 2018 for similar findings).

In another study, lie tellers and truth tellers were interviewed twice, immediately and two weeks after the event (Izotovas et al., 2018). Truth tellers showed a forgetting curve and lie tellers a stability bias; however, truth tellers' forgetting curve was not as pronounced as in Harvey et al.'s studies (2017) and truth tellers still reported more details than lie tellers in the delayed interview. Two differences between the studies of Harvey et al. and Izotovas et al. may have caused the differences in findings. First, in Izotovas et al., participants in the delay condition also had an early recall attempt, whereas such an attempt was absent in Harvey et al. The early recall attempt in the immediate interview in Izotovas et al. may have strengthened episodic memory, which in turn facilitated later recall by reducing the extent of forgetting (Gabbert, Hope, & Fisher, 2009; Gabbert, Hope, Fisher, & Jamieson, 2011). Moreover, unlike in Harvey et al., in the immediate interview in Izotovas et al., all participants were introduced to specific memory-enhancement tools (mnemonics including context reinstatement and sketches). This may have strengthened episodic memory (Dando, Wilcock, & Milne, 2009; Memon, Meissner, & Fraser, 2010).

## **2. The Current Experiment**

In the current experiment, participants were interviewed three times each time one week apart. The Model Statement was either introduced at Time 1 (MS Week 1 condition), Time 2



(MS Week 2 condition), Times 1 and 2 (MS Weeks 1 & 2 condition), or not at all (MS control condition). The Model Statement should encourage lie tellers and truth tellers to report more details, which would strengthen episodic memory of the event and thus facilitate later recollection. If participants listen to the Model Statement at Time 1, this facilitation pattern may continue at Times 2 and 3. If participants listen to the Model Statement at Time 2, this facilitation pattern may continue at Time 3. Perhaps most encouragement to provide information is given when the Model Statement is presented at Times 1 and 2. The type of detail truth tellers and lie tellers report after a Model Statement may also differ (more core details and complications by truth tellers and more peripheral details, common knowledge details and self-handicapping strategies by lie tellers).

### *2.1. Hypotheses*

The first set of hypotheses (Hypotheses 1 to 3) focused on Time 3 (Week 3) as we were interested in the effect of the Model Statement on the elicitation of information and cues to deception over time. The second set of hypotheses (Hypotheses 4 to 6) focused on unique details provided at each of Times 1, 2, and 3.

In line with previous findings on deception (e.g., Vrij, Leal, Mann, et al., 2018), it was expected that at Time 3, truth tellers would report more core details and complications than lie tellers, whereas lie tellers would report more peripheral details, common knowledge details, and self-handicapping strategies than truth tellers (Veracity main effect, *Hypothesis 1*).

The Model Statement was expected to strengthen memory (Vrij, Leal, Jupe, et al., 2018). At Time 3, most core details, peripheral details, complications, common knowledge details, and self-handicapping strategies would be elicited in the MS Weeks 1 & 2 condition, less so in the MS Week 1 and MS Week 2 conditions, and least in the MS control condition at Time 3 (Model Statement main effect, *Hypothesis 2*).

In line with Hypotheses 1 and 2, truth tellers and lie tellers at Time 3 would differ most from each other on core details, peripheral details, complications, common knowledge details and self-handicapping strategies in the MS Weeks 1 & 2 condition, less so in the MS Week 1 and MS Week 2 conditions, and least in the MS control condition at Time 3 (Veracity x Model Statement interaction, *Hypothesis 3*).

In real life multiple interview settings, one main aim of investigators is to obtain new information in each subsequent interview. Therefore, we examined the number of unique core, peripheral, and common knowledge details, unique self-handicapping strategies and unique complications across the three interviews. For each variable, we computed a score that included the corresponding type of detail at Time 1 + new details at Time 2 + new details at Time 3.

Truth tellers would report more unique core details and unique complications than lie tellers, whereas lie tellers would report more unique peripheral details, unique common knowledge details, and unique self-handicapping strategies than truth tellers (Veracity main effect, *Hypothesis 4*).

Most unique core details, peripheral details, complications, common knowledge details, and self-handicapping strategies would be elicited in the MS Weeks 1 & 2 condition, less so in the MS Week 1 and MS Week 2 conditions, and least in the MS control condition (Model Statement main effect, *Hypothesis 5*).

In line with Hypotheses 4 and 5, truth tellers and lie tellers would differ most from each other on unique core details, peripheral details, complications, common knowledge details and self-handicapping strategies in the MS Week 1 & 2 condition, less so in MS Week 1 and MS Week 2 conditions, and least in the MS control condition (Veracity x Model Statement interaction effect, *Hypothesis 6*).

## 2.2. Exploratory analyses

We carried out two sets of exploratory analyses that were not pre-registered. First, in recent years, researchers argued and found that the proportion of complications score discriminates truth tellers from lie tellers better than the generic variable total details (Vrij, Leal, Fisher, et al., 2018; Vrij, Leal, Fisher, Mann, Deeb, 2019; Vrij, Leal, Jupe, et al., 2018; Vrij, Leal, et al., 2017; Vrij, Leal, Mann, et al., 2018). The proportion of complications score takes into account the different types of detail truth tellers and lie tellers report (i.e., truth tellers report complications and lie tellers report common knowledge details and self-handicapping strategies), whereas the generic variable ‘total details’ does not. We sought to replicate this finding in the present experiment. Following the same reasoning, the proportion of core details defined as  $\text{core details} / (\text{core details} + \text{peripheral details})$  should also discriminate truth tellers from lie tellers better than the variable total details.

Second, researchers recommended the use of the Model Statement technique in real life as a within-subject tool so that it controls for individual and situational differences (Vrij, Leal, & Fisher, 2018). Therefore, the interview would start with an initial free recall, followed by a Model Statement, followed by a second free recall. New details reported in subsequent interviews are hereby the focus of attention. Although it is recommended that this within-subject approach is applied in a single interview, we wanted to explore its effects when introduced between interviews. We thus examined the change in participants’ reports from Time 1 to Time 2 to understand the incremental effects of the Model Statement on the dependent variables. The Time 3 data are not suitable for this analysis because none of the participants listened to the Model Statement in the third interview.

### **3. Method**

#### *3.1. Participants and design*

We recruited a total of 255 student and staff members at the University of Portsmouth but due to technical errors and the withdrawal of two participants, the final sample included 243 participants of whom 187 were females (77%) and 56 were males (23%),  $M_{age} = 23.89$ ,  $SD_{age} = 8.89$ . Approximately half (57%) of the participants were British, and the remaining participants were Asian (14%), European (12%), African (3%), Hispanic (3%), Arab (1%), and of mixed ethnicity (9%). One participant did not mention ethnicity. Although 43% of the participants were non-native English speakers, we do not anticipate that this would have affected the results as it has been shown that non-native but highly proficient speakers provide statements that are similar to those of native speakers (Evans, Pimentel, Pena, & Michael, 2017). As English is a requirement at the corresponding university, participants are expected to be highly proficient English speakers. To corroborate this conclusion, we examined potential differences between the British and non-British participants in our sample. We ran two one-way MANOVAs with core details, peripheral details, common knowledge details, self-handicapping strategies, complications, proportion of complications, proportion of core details and total details as dependent variables, and we found no significant differences for Time 3 data, Pillai's Trace = .02,  $F(7, 235) = 0.69$ ,  $p = .682$ ,  $\eta^2 = .02$ , or for unique details, Pillai's Trace = .01,  $F(7, 235) = 0.26$ ,  $p = .967$ ,  $\eta^2 = .01$ .

Participants either received two course credits or £30 for taking part in the study. They were also included in a draw to win one of three prizes (£50, £75, or £150). The study was approved by the standing ethics committee.

A 2 (Veracity: truth teller, lie teller) x 4 (Model Statement: MS Week 1, MS Week 2, MS Weeks 1 & 2, MS control) between-subjects design was conducted with core details, peripheral details, common knowledge details, self-handicapping strategies, complications, proportion of complications, proportion of core details and total details as dependent variables.

Participants were randomly allocated to the Veracity and Model Statement conditions. The sample included 120 truth tellers and 123 lie tellers. Among truth tellers, 30 participants were allocated to the MS Week 1 condition, 32 to the MS Week 2 condition, 29 to the MS Weeks 1 & 2 condition, and 29 to the MS control condition. Among lie tellers, 31 were allocated to the MS Week 1 condition, 31 to the MS Week 2 condition, 31 to the MS Weeks 1 & 2 condition, and 30 to the MS control condition.

### *3.2. Procedure*

#### *3.2.1. Initial instructions*

The procedure was adapted from Leal et al. (2018). Upon arrival to the psychology department (Time 1), all participants signed a consent form. Truth tellers were asked to think of a memorable event that happened to them in the past two years. The event should have been an out of the ordinary event that they (or others) do not experience daily. Truth tellers were given examples of such events (e.g., seeing a famous person when they were out for dinner, or going on a pheasant shoot). Lie tellers and truth tellers were matched for memorable events. That is, lie tellers were asked to lie about an event previously chosen by a truth teller and to pretend they experienced that event. Matched lie tellers and truth tellers were allocated to the same Model Statement condition.

All participants were informed that they will be interviewed about the event (in three interviews, each one week apart) and they will need to convince the interviewer they are honest. To motivate participants, they were told that if they are convincing, their names will be entered in a draw to win £50, £75, or £150. However, if they are not convincing, they will be asked to write a statement about the event and their names will not be entered in the draw. At Times 2 and 3, participants were given the same motivating instructions again and told that their name will be entered in the draw for a second (and/or third) time if they are convincing in the corresponding interview; otherwise, if they are not convincing, they will

have to write a statement. In reality, all participants were entered in the draw and none of them was asked to write a statement. All participants were given as much time as they need to prepare for the interview.

### *3.2.2. Pre-interview questionnaires*

Before each interview, participants completed an automated pre-interview questionnaire that was created via Qualtrics software. Participants rated (a) their motivation to appear convincing during the interview, and (b) the extent to which their preparation for the interview was sufficient, good, and thorough (1 = *not at all* to 7 = *completely*). They were also asked whether or not they prepared any strategies for the interview and to write down the preparation strategy or the reason for not preparing a strategy. In addition, at Time 1, participants were asked about their background characteristics, and at Times 2 and 3, they were asked whether or not (and for how long) they prepared for the interview during the past week outside the interview setting.

### *3.2.3. Interviews*

After completing the pre-interview questionnaire, participants were taken to the interview room where they were interviewed by one of three female research assistants (interviewers), blind to the participants' veracity condition and to the study hypotheses. The same interviewer interviewed each participant at Times 1, 2, and 3. However, in some cases and for reasons beyond our control, the same interviewer could not be available, so a different interviewer took this role. A multivariate analysis was conducted to examine the effect of having the same or a different interviewer across interviews on core, peripheral, and common knowledge details, complications, and self-handicapping strategies at Time 3. No significant differences emerged, Pillai's Trace = .04,  $F(10, 474) = .94$ ,  $p = .500$ ,  $\eta^2 = .02$ . Similarly, having the same or a different interviewer across interviews did not affect unique core,

peripheral, and common knowledge details, complications, and self-handicapping strategies, Pillai's Trace = .06,  $F(10, 474) = 1.40$ ,  $p = .175$ ,  $\eta^2 = .03$ . Thus, changing interviewers did not have any effect on the results.

In all three interviews, participants were asked only one question to report the event in as many details as possible. Participants who were not allocated to listen to the Model Statement were asked: "I understand from my colleague [the experimenter] that on [date of event], you [event]. Could you please tell me in as much detail as possible everything about the event?" If participants were allocated to listen to the Model Statement, the audiotape was played at the beginning of the interview before participants provided their responses. The following instructions were given:

I would like you to tell me in as much detail as possible everything about that event, but before you do so I am going to play you a Model Statement in order to give you an idea of exactly how much detail I would like you to provide.

The Model Statement is a 1.30 minutes long detailed account of someone attending a Formula 2 motor racing event (see Leal et al., 2015). The account is a spontaneous, unscripted recall of the event that is truly experienced by the person. After listening to the Model Statement, participants were instructed: "Remembering the amount of detail in the Model Statement, could you please tell me what happened on [date of event] when you [event]?"

Participants allocated to the MS Week 1 condition listened to the Model Statement at the beginning of the interview at Time 1, those allocated to the MS Week 2 condition listened to the Model Statement at the beginning of the interview at Time 2, those allocated to the MS Weeks 1 & 2 condition listened to the Model Statement at Times 1 and 2, and those in the MS control condition did not listen to the Model Statement at all.

#### *3.2.4. Post-interview questionnaire*

After the interview, participants were either scheduled for the next session (Times 1 and 2) or were asked to complete an automated post-interview questionnaire (at Time 3) that was created via Qualtrics software. They rated on 7-pt scales (1= *not at all* to 7 = *completely*) their motivation, the extent to which they thought— in the past two weeks outside the interview setting—of the event or of what they were going to say, perceived believability by the interviewer, likelihood of writing a statement, probability of winning the prize, interviewer's characteristics (smooth, bored, satisfied, awkward, engrossed, involved, friendly, active, and positive), willingness to volunteer information in the three interviews, and the extent to which they limited their report to discussing core details in order to shorten the interview. On percentage scales (0-100%), they rated the extent to which they disclosed or concealed information they remembered and the extent to which they were truthful and deceptive in the three interviews. As a manipulation check, participants were asked if their event was genuine or made up. In addition, participants were asked about their preparation strategy, if they used a different strategy than originally planned, and if they previously participated in Model Statement studies. Also, participants who listened to the Model Statement, were asked if they did so carefully, if they were thinking of what to say while listening to the Model Statement, and if the Model Statement prompted them to provide fewer or more details than originally planned (1= *not at all* to 7 = *completely*). After completing the questionnaire, participants were fully debriefed, rewarded, and thanked.

### 3.3. Coding

All interviews were audiotaped, transcribed and coded for details. Every noun, verb, adjective, and adverb was considered a single detail, but other details such as conjunctions, prepositions, and pronouns were not counted because they are not precise and thus not informative. For example, the statement "My three friends went to a restaurant and drank cocktails" includes six details (three, friends, went, restaurant, drank, cocktails). A distinction



was made between core and peripheral details. To illustrate how we coded core and peripheral details, we use an example of a participant who talked about a lost challenge resulting in eating a scorpion during a visit to Bangkok. The participant's statement "We lost the dominoes game 200 to 150 so we had to eat a scorpion" describes information that concerned the target event and includes seven core details (lost, dominoes, game, 200, 150, eat, scorpion). The same participant also provided information that was peripheral or irrelevant to the target event: "We went to a hotel that a friend recommended". This statement includes four peripheral details (went, hotel, friend, recommended).

Details that were repeated in a single interview were coded only once. We also coded new details at Times 2 and 3. Details mentioned in the second or third recall but not in a previous recall were coded as new details, and classified into core or peripheral details.

The first author and another coder—a research assistant who was not involved in any other aspect of the research—coded the transcripts independently for core and peripheral details. Both coders were blind to participants' veracity conditions. They started by coding interviews from two participants. Disagreements between them were discussed and resolved. The first author coded all the transcripts, and the other coder coded 15% of the transcripts. Inter-rater reliability analyses were computed using the Intra-Class Correlation (ICC) coefficient (single measures scores). Hallgren (2012) reports that inter-rater reliability is poor for ICC values less than .40, fair for values between .40 and .59, good for values between .60 and .74, and excellent for values between .75 and 1. ICC was excellent for core details (ICC = 0.87) and good for peripheral details (ICC = 0.61).

Further, the second author and a research assistant, blind to participants' veracity conditions, coded the transcripts independently for common knowledge details (strongly invoked stereotypical knowledge about events), self-handicapping strategies (explicit or implicit justifications as to why someone is not able to provide information), and

complications (occurrences that make a situation more difficult than necessary). The statement: “I had loads of fun spending time with my friends” contains one common knowledge detail. The statement “It was in the evening so it was quite dark, so I didn’t see a lot” contains one self-handicapping strategy. The statement (about a hockey game): “Then one of the girls got a shoulder injury and I realised that I was the next sub, and I didn’t bring my own equipment. I had to borrow hers which was all smelly and sweaty.” includes three complications (shoulder injury, did not bring own equipment, having to borrow hers). One coder coded 32% of the transcripts and the other coder coded all the transcripts. Inter-rater reliability was good for common knowledge details (ICC = 0.60), self-handicapping strategies (ICC = 0.70), and complications (ICC = 0.72).

Based on the coded data, the proportion scores of (i) core details and (ii) complications and the total number of details were calculated. The proportion of core details (at Time 3) was calculated as:  $\text{core details Time 3} / (\text{core details Time 3} + \text{peripheral details Time 3})$ . The proportion of complications (at Time 3) was calculated as:  $\text{complications Time 3} / (\text{complications Time 3} + \text{common knowledge details Time 3} + \text{self-handicapping strategies Time 3})$ . The total number of details (at Time 3) included the total number of core and peripheral details (at Time 3).

Unique details were details mentioned at Time 1 added to new details mentioned at Time 2 and Time 3. The proportion of unique core details was calculated by dividing unique core details across the three interviews by the total number of unique core and peripheral details across interviews. The proportion of unique complications was the number of unique complications across the three interviews divided by the number of unique complications, common knowledge details, and self-handicapping strategies across interviews. Unique total details were total details at Time 1 added to new details at Time 2 and Time 3.

Participants' open responses in the pre-interview and post-interview questionnaires regarding strategies used and reasons for (not) using a strategy or a different strategy than originally planned were coded independently by the first author and a research assistant (also not involved in any other aspect of the research), both blind to the veracity conditions. Categories were formulated based on participants' responses to the questions. That is, similar responses were grouped together in a single category (theme), and each category was named to describe the theme of responses. When the same response could fit in more than one category, it was allocated to those corresponding categories. The categories representing participants' responses and their respective frequencies are reported in Table 1. Inter-rater agreement was substantial (Hallgren, 2012), Cohen's  $\kappa = 0.62$ . Although the agreement falls on the lower bound of the substantial range, this is not uncommon in deception research (Vrij, Leal, Mann, & Fisher, 2012). Nonetheless, disagreements were discussed and resolved.

Table 1

*Frequency of Convincing Strategies Reported by Truth tellers and Lie tellers in the Post-Interview Questionnaire*

<b>Convincing strategies</b>	<b>Truth tellers (<i>n</i> = 39)</b>	<b>Lie tellers (<i>n</i> = 84)</b>
Be detailed	8	14
Monitor non-verbal behaviour	16	34
Rely on memory/Be truthful	13	0
Include verifiable details	3	6
Rehearse for the interview	5	17
Be consistent	6	17
Report events chronologically	4	3
Be simple	4	9
Use embedded lies	0	21
Be inconsistent	0	7
Be spontaneous	2	7

*Note.* The total number in each column may exceed the number of truth tellers/ lie tellers in that column because the same participant may have reported more than one strategy.

#### 4. Results

A power analysis via G\*Power software revealed that perfect power was achieved ( $\alpha = 0.05, f^2 = 0.25$ ). The detailed results of all pre- and post-interview questionnaires with descriptive and inferential statistics are included in the Appendices.

##### 4.1. Testing hypotheses 1 through 3: Time 3 data

A multivariate analysis of variance (MANOVA) was used to control for Type I error as we tested several dependent variables. A 2 (Veracity: lie teller, truth teller) x 4 (Model Statement: Week 1, Week 2, Weeks 1 & 2, control) MANOVA with core details, peripheral details, common knowledge details, self-handicapping strategies and complications at Time 3 as dependent variables, revealed significant multivariate effects of Veracity, Pillai's Trace = .23,  $F(5, 231) = 13.76, p < .001, \eta^2 = .23$ , and Model Statement, Pillai's Trace = .17,  $F(15, 699) = 2.83, p < .001, \eta^2 = .06$ . The Veracity x Model Statement interaction effect was not significant, Pillai's Trace = .03,  $F(15, 699) = 0.53, p = .924, \eta^2 = .01$ . We report below the results of the univariate analyses along with Hedge's  $g$  which is the unbiased correction of Cohen's  $d$  effect size, Cohen's  $U_3$  which presents the non-overlap between means, and Probability of Superiority ( $PS$ ) which demonstrates the probability that a randomly selected person from one group has a higher observed score than a randomly selected person from the other group (Fritz, Morris, & Richler, 2012; Lakens, 2013; also see <https://rpsychologist.com/d3/cohend/>).

Table 2 shows that truth tellers reported more core details and complications than lie tellers. In contrast, lie tellers reported more common knowledge details and self-handicapping strategies than truth tellers. Therefore, Hypothesis 1 that predicted a Veracity main effect for all dependent variables was supported for all variables except for peripheral details.

As for the Model Statement main effect, the MS Week 1 condition elicited more common knowledge details ( $M = 1.79$ ,  $SD = 2.43$ ) than the MS control condition ( $M = 0.69$ ,  $SD = 1.47$ , Hedge's  $g = 0.54$ , 95% CI [0.18, 0.91], Cohen's  $U_3 = 70.54$ ,  $PS = 64.87$ ) and the MS Weeks 1 & 2 condition ( $M = 0.65$ ,  $SD = 0.97$ , Hedge's  $g = 0.61$ , 95% CI [0.25, 0.98], Cohen's  $U_3 = 72.91$ ,  $PS = 66.69$ ),  $F(3, 235) = 6.22$ ,  $p < .001$ ,  $\eta^2 = .07$ . Thus, Hypothesis 2 that predicted a Model Statement main effect was supported for common knowledge details only. As we did not find a Model Statement x Veracity effect, Hypothesis 3 that predicted an interaction effect was not supported.

Table 2

*Descriptive and Inferential Statistics of Time 3 Data as a Function of Veracity*

Detail type	Truth tellers $n = 120$ $M (SD)$	Lie tellers $n = 123$ $M (SD)$	$F$	$p$	Hedge's $g$ [95% CI]	Cohen's $U_3$	$PS$	$BF_{10}$
Core details	65.06 (37.00)	40.27 (27.48)	36.36	<.001	0.76 [0.50, 1.02]	77.64	70.45	1.026 x 10 <sup>6</sup>
Peripheral details	10.26 (16.36)	14.53 (22.15)	02.90	.090	0.22 [-0.03, 0.47]	58.71	56.18	0.554
Common knowledge details	00.73 (01.55)	01.43 (01.84)	11.20	.001	0.41 [0.16, 0.66]	65.91	61.41	18.279
Self-handicapping strategies	00.00 (00.00)	00.11 (00.43)	08.33	.004	0.36 [0.11, 0.62]	64.06	60.05	7.337
Complications	05.93 (04.82)	03.23 (04.13)	21.83	<.001	0.60 [0.34, 0.86]	72.57	66.43	3302.616
Proportion of core details	00.86 (00.16)	00.78 (00.20)	13.84	<.001	0.44 [0.19, 0.70]	67.00	62.21	96.733
Proportion of complications	00.84 (00.28)	00.57 (00.40)	35.75	<.001	0.78 [0.52, 1.04]	78.23	70.94	1.624 x 10 <sup>6</sup>
Total details	75.32 (40.84)	54.80 (40.33)	16.13	<.001	0.50 [0.25, 0.76]	69.15	63.82	184.283

*Note.*  $M$  = Mean;  $SD$  = Standard deviation;  $PS$  = Probability of Superiority.

We ran an exploratory 2 (Veracity: lie teller, truth teller) x 4 (Model Statement: Week 1, Week 2, Weeks 1 & 2, control) MANOVA with proportion of core details, proportion of complications, and total details at Time 3 as dependent variables. A significant multivariate effect of Veracity emerged, Pillai's Trace = .22,  $F(3, 233) = 21.90$ ,  $p < .001$ ,  $\eta^2 = .22$ . The Model Statement main effect, Pillai's Trace = .05,  $F(9, 705) = 1.36$ ,  $p = .203$ ,  $\eta^2 = .02$ , and

the Veracity x Model Statement interaction effect, Pillai's Trace = .02,  $F(9, 705) = 0.60$ ,  $p = .799$ ,  $\eta^2 = .01$ , were not significant. As Table 2 shows, the proportion scores of core details and complications and the total details were higher among truth tellers than lie tellers.

To further test our results, separate Bayesian ANOVAs with Model Statement and Veracity as factors were conducted via JASP 9.2 Software to compare evidence that at Time 3, truth tellers were more likely than lie tellers to report (proportion of) core details and complications, and total details and less likely to report peripheral details, common knowledge details, and self-handicapping strategies (H1) with evidence that this is unlikely (H0). Bayes factors ( $BF_{10}$ ) that are between 1 and 3 indicate weak evidence for H1, between 3 and 20 indicate positive evidence, between 20 and 150 indicate strong evidence, and above 150 indicate very strong evidence (Jarosz & Wiley, 2014; Raftery, 1995).

Bayes factors ( $BF_{10}$ ) for the Veracity main effect are reported in Table 2. Positive to very strong evidence emerged for all variables except for peripheral details which received weak evidence. Also, evidence for the proportion of complications was stronger than evidence for the proportion of core details and total details. For the Model Statement main effect, weak evidence emerged for all variables except for common knowledge details which received strong evidence ( $BF_{10} = 35.911$ ).

Moreover, the Veracity x Model Statement interaction effect gained positive to very strong support for some of the dependent variables. However, the Bayesian inclusion analyses revealed that the data was supported only when Veracity was included in the model (all  $BF_{Inclusion} < 1$  for the interaction effect), meaning that evidence supported an effect for Veracity only. These results replicated the frequentist analyses (significance testing) results reported previously.

#### 4.2. Testing hypotheses 4 through 6: Unique details

A 2 (Veracity: lie teller, truth teller) x 4 (Model Statement: Week 1, Week 2, Weeks 1 & 2, control) MANOVA with unique core, peripheral, and common knowledge details, self-handicapping strategies, and complications as dependent variables revealed significant multivariate effects of Veracity, Pillai's Trace = .23,  $F(5, 231) = 13.93$ ,  $p < .001$ ,  $\eta^2 = .23$ , and Model Statement, Pillai's Trace = .13,  $F(15, 699) = 2.15$ ,  $p = .007$ ,  $\eta^2 = .04$ . The Veracity x Model Statement interaction effect was not significant, Pillai's Trace = .04,  $F(15, 699) = 0.56$ ,  $p = .905$ ,  $\eta^2 = .01$ . Table 3 shows that truth tellers included more unique core details and complications and fewer unique peripheral details, common knowledge details, and self-handicapping strategies than lie tellers.

Table 3

*Descriptive and Inferential Statistics of Unique Data as a Function of Veracity*

Detail type	Truth tellers <i>n</i> = 120 <i>M</i> ( <i>SD</i> )	Lie tellers <i>n</i> = 123 <i>M</i> ( <i>SD</i> )	<i>F</i>	<i>P</i>	Hedge's <i>g</i> [95% CI]	Cohen's <i>U</i> <sub>3</sub>	<i>PS</i>	<i>BF</i> <sub>10</sub>
Unique core details	100.90 (60.98)	65.25 (40.42)	29.52	<.001	0.69 [0.43, 0.95]	75.49	68.72	69669.113
Unique peripheral details	18.39 (23.81)	27.44 (41.73)	4.21	.041	0.26 [0.01, 0.52]	60.26	57.29	1.052
Unique common knowledge details	01.44 (02.27)	02.50 (02.92)	10.52	.001	0.40 [0.15, 0.66]	65.54	61.14	14.873
Unique self-handicapping strategies	00.06 (00.24)	00.29 (00.73)	11.04	.001	0.42 [0.17, 0.68]	66.28	61.68	25.284
Unique complications	09.16 (07.58)	04.92 (06.29)	22.21	<.001	0.61 [0.35, 0.87]	72.91	66.69	4220.273
Proportion of unique core details	00.84 (00.16)	00.75 (00.21)	14.45	<.001	0.48 [0.23, 0.74]	68.44	63.29	127.753
Proportion of unique complications	00.84 (00.24)	00.59 (00.35)	40.73	<.001	0.83 [0.57, 1.09]	79.67	72.14	1.215 x 10 <sup>7</sup>
Total unique details	119.29 (67.00)	92.69 (66.24)	10.03	.002	0.40 [0.14, 0.65]	65.54	61.14	12.875

*Note.* *M* = Mean; *SD* = Standard deviation; *PS* = Probability of Superiority.

As for the Model Statement main effect, a significant univariate effect emerged for unique common knowledge details,  $F(3, 235) = 4.66$ ,  $p = .003$ ,  $\eta^2 = .06$ . The MS Week 1 condition elicited more unique common knowledge details ( $M = 2.67$ ,  $SD = 3.34$ ) than the MS Weeks 1&2 condition ( $M = 1.42$ ,  $SD = 1.68$ , Hedge's  $g = 0.47$ , 95% CI [0.11, 0.83],

Cohen's  $U_3 = 68.08$ ,  $PS = 63.02$ ) and the MS control condition ( $M = 1.31$ ,  $SD = 2.14$ , Hedge's  $g = 0.48$ , 95% CI [0.12, 0.85], Cohen's  $U_3 = 68.44$ ,  $PS = 63.29$ ).

An exploratory 2 (Veracity: lie teller, truth teller) x 4 (Model Statement: Week 1, Week 2, Weeks 1 & 2, control) MANOVA was conducted with the proportion of unique core details, proportion of unique complications, and total unique details as dependent variables. The analysis revealed a significant multivariate effect of Veracity, Pillai's Trace = .21,  $F(3, 233) = 20.51$ ,  $p < .001$ ,  $\eta^2 = .21$ . The Model Statement main effect, Pillai's Trace = .04,  $F(9, 705) = 1.05$ ,  $p = .402$ ,  $\eta^2 = .01$ , and the Veracity x Model Statement interaction effect, Pillai's Trace = .02,  $F(9, 705) = 0.55$ ,  $p = .837$ ,  $\eta^2 = .01$ , were not significant. As Table 4 shows, all three variables were higher among truth tellers than lie tellers.

Table 4

*Descriptive and Inferential Statistics of Time 2 Data as a Function of Veracity*

Detail type	Truth tellers $n = 120$ $M (SD)$	Lie tellers $n = 123$ $M (SD)$	$F$	$P$	Hedge's $g$ [95% CI]	Cohen's $U_3$	$PS$	$BF_{10}$
New core details	19.16 (17.96)	14.15 (14.44)	06.69	.010	0.31 [0.05, 0.56]	62.17	58.68	2.101
New peripheral details	04.77 (09.74)	08.43 (16.52)	04.61	.033	0.27 [0.02, 0.52]	60.64	57.57	1.117
New common knowledge details	00.28 (00.76)	00.59 (01.44)	05.07	.025	0.27 [0.02, 0.52]	60.64	57.57	1.248
New self-handicapping strategies	00.02 (00.13)	00.09 (00.36)	04.44	.036	0.26 [0.00, 0.51]	60.26	57.29	1.066
New complications	02.13 (02.96)	00.93 (01.45)	17.22	<.001	0.52 [0.26, 0.77]	69.85	64.34	243.330
Proportion of new core details	00.82 (00.21)	00.69 (00.29)	15.85	<.001	0.51 [0.26, 0.77]	69.50	64.08	227.014
Proportion of new complications	00.58 (00.47)	00.40 (00.46)	09.84	.002	0.39 [0.13, 0.64]	65.17	60.86	11.592
Total new details	23.93 (23.59)	22.58 (23.17)	00.18	.670	0.06 [-0.19, 0.31]	52.39	51.69	0.155

*Note.*  $M$  = Mean;  $SD$  = Standard deviation;  $PS$  = Probability of Superiority;  $BF_{10}$  = Bayes Factor.

Separate Bayesian ANOVAs with Veracity and Model Statement as factors revealed a Veracity main effect—as shown in Table 3—with positive to very strong evidence for all dependent variables except for unique peripheral details, which received weak evidence. The proportion of unique core details and the proportion of unique complications received



stronger evidence than total unique details. As for the Model Statement main effect, positive evidence emerged for unique common knowledge details ( $BF_{10} = 5.147$ ) and weak evidence for all other dependent variables.

The Model Statement x Veracity interaction effect gained positive to very strong support, but the Bayesian inclusion factors for the Model Statement were below 1, and thus only the Veracity effect was supported. These results replicated the frequentist analyses (significance testing results) reported above—except for unique peripheral details. Overall, Hypothesis 4 that predicted a Veracity main effect was supported for all variables except for unique peripheral detail. Hypothesis 5 that predicted a Model Statement main effect for all dependent variables was supported for unique common knowledge details only. Hypothesis 6 that predicted an interaction effect was not supported.

The results of the unique data analyses were virtually identical to those of the Time 3 data analyses. The only difference was that a significant effect for peripheral details emerged in the unique data analyses but not in the Time 3 data analyses. However, this effect was weak according to the Bayes Factors analyses.

#### *4.3. Further exploratory analyses: Time 2 data*

We conducted further exploratory analyses on the Time 2 new details. A 2 (Veracity: lie teller, truth teller) x 4 (Model Statement: Week 1, Week 2, Weeks 1 & 2, control) MANOVA was carried out with new core details, peripheral details, common knowledge details, self-handicapping strategies and complications at Time 2 as dependent variables. Significant multivariate effects emerged for Veracity, Pillai's Trace = .16,  $F(5, 231) = 8.63$ ,  $p < .001$ ,  $\eta^2 = .16$ , and Model Statement, Pillai's Trace = .28,  $F(15, 699) = 4.81$ ,  $p < .001$ ,  $\eta^2 = .09$ . The Veracity x Model Statement interaction effect was not significant, Pillai's Trace = .09,  $F(15, 699) = 1.49$ ,  $p = .103$ ,  $\eta^2 = .03$ .

As Table 4 shows, truth tellers reported significantly more new core details and complications and fewer new peripheral details, common knowledge details, and self-handicapping strategies than lie tellers. Bayesian ANOVAs revealed that all these Veracity effects received weak evidence, except the effect for new complications. As for the Model Statement effects (see Table 5), participants in the MS Week 2 condition provided significantly more new core details and complications than participants in the other conditions (MS Week 1, MS Weeks 1 & 2, and MS control conditions). The MS Weeks 1 & 2 condition elicited significantly more new core details than the MS Week 1 and MS control conditions.

Table 5

*Descriptive and Inferential Statistics of Time 2 Data as a Function of Model Statement Condition*

Detail type	MS 1 <i>n</i> = 61 <i>M</i> ( <i>SD</i> )	MS 2 <i>n</i> = 63 <i>M</i> ( <i>SD</i> )	MS 1&2 <i>n</i> = 60 <i>M</i> ( <i>SD</i> )	No MS <i>n</i> = 59 <i>M</i> ( <i>SD</i> )	<i>F</i>	<i>p</i>	$\eta^2$	<i>BF</i> <sub>10</sub>
New core details	10.92 <sup>a</sup> (10.49)	27.43 <sup>c</sup> (19.14)	17.93 <sup>b</sup> (17.04)	09.64 <sup>a</sup> (10.64)	18.98	<.001	.20	1.274×10 <sup>8</sup>
New peripheral details	06.07 (12.84)	10.67 (19.78)	05.95 (10.64)	03.56 (07.11)	3.06	.029	.04	0.827
New common knowledge details	00.48 <sup>ab</sup> (01.13)	00.89 <sup>b</sup> (01.83)	00.17 <sup>a</sup> (00.42)	00.19 <sup>a</sup> (00.39)	5.83	.001	.07	19.978
New self-handicapping strategies	00.02 (00.13)	00.13 (00.42)	00.03 (00.26)	00.03 (00.18)	2.19	.090	.03	0.280
New complications	00.84 <sup>a</sup> (01.25)	02.78 <sup>b</sup> (03.67)	01.37 <sup>a</sup> (01.65)	01.03 <sup>a</sup> (01.64)	9.71	<.001	.11	1944.737
Proportion of new core details	00.74 (00.27)	00.78 (00.25)	00.77 (00.26)	00.74 (00.27)	0.52	.668	.01	0.038
Proportion of new complications	00.37 <sup>a</sup> (00.47)	00.62 <sup>b</sup> (00.43)	00.57 <sup>ab</sup> (00.48)	00.40 <sup>ab</sup> (00.47)	4.54	.004	.06	5.110
Total new details	16.98 <sup>ab</sup> (19.77)	38.10 <sup>c</sup> (28.02)	23.88 <sup>b</sup> (21.12)	13.20 <sup>a</sup> (13.84)	16.32	<.001	.17	8.550×10 <sup>6</sup>

*Note.* Only mean scores with a different superscript differ significantly from each other ( $p < .05$ ). *M* = Mean; *SD* = Standard deviation; *BF*<sub>10</sub> = Bayes Factor; MS 1 = MS Week 1 condition; MS 2 = MS Week 2 condition; MS 1&2 = MS Weeks 1 & 2 condition; No MS = MS control condition.

The MS Week 2 condition elicited more new common knowledge details than in the MS Weeks 1 & 2 and the MS control condition. The MS Week 2 condition also elicited more new peripheral details than in the MS control condition. Bayesian analyses revealed strong to

very strong evidence for these effects, except for new peripheral details which was weakly supported.

A 2 (Veracity: lie teller, truth teller) x 4 (Model Statement: Week 1, Week 2, Weeks 1 & 2, control) mixed MANOVA was conducted on the proportion of core details, proportion of complications, and total details at Time 2. Significant multivariate effects emerged for Veracity, Pillai's Trace = .10,  $F(3, 233) = 8.58$ ,  $p < .001$ ,  $\eta^2 = .10$ , and Model Statement, Pillai's Trace = .21,  $F(9, 705) = 5.94$ ,  $p < .001$ ,  $\eta^2 = .07$ . The Veracity x Model Statement interaction effect was not significant, Pillai's Trace = .04,  $F(9, 705) = 0.98$ ,  $p = .452$ ,  $\eta^2 = .01$ . As Table 4 shows, the proportion scores for new core details and complications were significantly higher for truth tellers than lie tellers. Total details did not differ between lie tellers and truth tellers.

As for the Model Statement main effect (see Table 5), the proportion score of complications was significantly higher in the MS Week 2 condition than in the MS Week 1 condition. Also, total details were significantly higher in the MS Week 2 condition than in all other conditions, and in the MS Weeks 1 & 2 condition than in the MS control condition.

## 5. Discussion

In the current research, we attempted to understand the effects of using the Model Statement on deception detection and information elicitation when the same interviewee is interviewed multiple times. Our hypotheses focused on the information provided in the third interview and on unique information provided during the three interviews. The analyses showed many Veracity main effects but limited Model Statement main effects. No Veracity x Model Statement interaction effects occurred.

### *5.1. The effects of multiple interviews on the reports of lie tellers and truth tellers*

The results showed that in a third interview, truth tellers reported more core and total details and complications than lie tellers, whereas lie tellers reported more common knowledge details and self-handicapping strategies. Truth tellers also reported higher proportions of core details and complications than lie tellers. The same pattern of results emerged for unique details (details mentioned at Time 1 and new details mentioned at Time 2 or Time 3). This overlap in results means that if, for whatever reason, investigators only have access to the Time 3 interviews, they would not be in a disadvantageous position for assessing veracity compared to investigators who have access to all three interviews.

Our findings align with previous findings on the Model Statement within single interviews which showed differences between lie tellers and truth tellers in terms of core details, total details, complications, common knowledge details, and self-handicapping strategies (Leal et al., 2018; Vrij, Leal, Jupe, et al., 2018). Our results also replicated previous findings that lie tellers and truth tellers differ on core details (Roos af Hjelmsäter, Öhman, Granhag, & Vrij, 2014; Sakrisvold et al., 2017), total details (see Amado et al. [2016] for a meta-analysis), complications, common knowledge details, self-handicapping strategies and proportion of complications (Vrij, Leal, Jupe, & Harvey, 2018; Vrij et al., 2017, 2018) within single interviews. Apparently, our findings can be generalised to different interview settings, adding to the robustness of such cues in assessing veracity.

It is reassuring that strong Veracity effects still occurred at Time 3 and when examining unique details in all three interviews, given that we used a very basic interview protocol (asking just one open ended question) and repeated that basic interview protocol three times. In other words, we did not try to catch out lie tellers through difficult questioning. The self-reported strategies (Table 1) give insight into the possible countermeasures lie tellers used and why they did not yield success. By far the most frequently reported strategy was monitoring nonverbal behaviour. Using that strategy would not have affected the verbal cues

we were interested in. The three verbal strategies most frequently mentioned were using embedded lies, rehearsing stories, and being consistent. Rehearsing stories may be related to being consistent: Someone probably rehearses a story to maintain consistency. Since we did not examine consistency, using the rehearsing stories and being consistent strategies would not have affected our results either. It is possible that the embedded lies strategy assisted lie tellers in providing truthful information and sounding like truth tellers. However, evidence from the pre-interview and post-interview questionnaires (see Tables A.1 and B.1 in the Appendices) and from previous deception research (Jundi, Vrij, Hope, Mann, & Hillman, 2013; Vrij, Mann, Jundi, Hillman, & Hope, 2014) suggests that even when lie tellers report embedded lies, they prefer to keep their stories simple.

We ran exploratory analyses to compare the proportion scores of (i) complications and (ii) core details with the total number of details. The cue ‘proportion score of complications’ was examined previously, but our study is the first to examine the cue ‘proportion of core details’. We found that the proportion of complications had the largest Veracity effect compared to the other two scores at Time 3 and when unique details were considered. This result replicates previous research that found that the proportion of complications was a more diagnostic indicator of veracity than total details (Vrij, Leal, Fisher, Mann, Jo, et al., 2019; Vrij, Leal, Jupe, et al., 2018; Vrij, Leal, Mann, et al., 2018; Vrij, Leal, et al., 2017). It also further demonstrates the benefit of examining a mixture of cues to truthfulness (complications) and cues to deception (common knowledge details and self-handicapping) versus just examining cues to truthfulness (total details).

The effect sizes for proportion of core details and total details were somewhat similar when the Time 3 details (Table 2) and unique details (Table 3) were analysed. Truth tellers and lie tellers mainly focused on the core event, which decreased the likelihood that a stronger Veracity effect would occur for proportion of core details than for total details. This

may be the result of the instructions given to truth tellers and lie tellers (to describe a memorable event), because the memorable event is by definition a core event. Maybe a more ambiguous scenario (e.g., asking interviewees to describe a particular day in its entirety, whilst the interviewer is actually only interested in a specific time slot) would yield different results.

We explored new information at Time 2 to examine the effects of multiple interviews on eliciting new details. The pattern of Veracity main effects for Time 2 new details was similar to that of Time 3 details and unique details; however, the Bayesian analyses showed that the effects for the Time 2 results were weak, except for new complications, which revealed very strong results. Therefore, for lie detection purposes, it may be more beneficial to examine information provided at Time 3 or unique details across interviews than it is to examine new information reported at Time 2.

### *5.2. The effects of presenting the Model Statement in multiple interviews*

The Model Statement main effects were identical for Time 3 data and unique details. In both analyses, participants in the MS Week 1 condition provided more common knowledge details than participants in the MS Weeks 1 & 2 condition and the MS control condition. We can only speculate why this effect occurred. A Model Statement typically raises expectations amongst interviewees regarding the amount of detail they are expected to report. Participants who listened to the Model Statement at Time 1 (Week 1) may have responded by providing more common knowledge details to fill in gaps in their reports. However, listening to the Model Statement again at Time 2 (MS Weeks 1 & 2) would have cautioned participants against providing common knowledge details. They may have thought that, because they were played the Model Statement again, the interviewer was not satisfied with their response to the Model Statement in the first interview and opted not to report them any longer.

When we analysed Time 2 results (Table 5), we found that for the purpose of eliciting new information, presenting a Model Statement at Time 2 is particularly effective.

Participants in the MS Week 2 condition provided more new core details, new complications, and new total details than those in the MS Week 1, MS Weeks 1 & 2, and MS control conditions. Participants in the MS Week 2 condition also provided more new common knowledge details than those in the MS Weeks 1 & 2 and MS control conditions, and had a higher proportion score of new complications than those in the MS Week 1 condition. The MS Weeks 1 & 2 condition was somewhat effective as it elicited more new core details than the MS Week 1 and MS control conditions and more new total details than the MS control condition.

The overall results suggest that introducing the Model Statement in a first interview can have carry over effects in eliciting common knowledge details; however, it is unlikely to elicit new details at a later stage. It is more likely to elicit new details when introduced in a subsequent interview.

### *5.3. Explanation for the lack of an interaction effect*

We did not find a Veracity x Model Statement interaction effect in any of the analyses despite predicting that based on theory. It might be that there was a ceiling effect. The Veracity effects for the Time 3 and unique details analyses were already very strong in the no Model Statement control condition, which gives little room for improvement in the Model Statement conditions. The strong Veracity effects across the Model Statement conditions may have been caused by the chosen deception scenario. The reported events experienced by truth tellers were not ordinary events that they would experience daily so they could add idiosyncratic details to their stories. In contrast, lie tellers were instructed to pretend to have experienced such unique events and may have had difficulty doing this due to the restrictive nature of the event. Although lie tellers used embedded lies, the results seemed to indicate

that the truthful aspects of their lies included common knowledge details (e.g., experiences known to them through acquaintances, TV programs, etc.) rather than idiosyncratic features. This suggests that different findings may emerge if truth tellers report less unique events and/or if lie tellers are given broad instructions to report events that are not unique.

A second explanation is that we did not introduce the Model Statement in the way recommended in the literature (Vrij, Leal, & Fisher, 2018). That is, to start an interview with a free recall, followed by a Model Statement, followed by another free recall. Whether such a format would yield Veracity x Model Statement interaction effects in multiple interviews is worth investigating.

#### *5.4. Methodological considerations*

Participants were asked to report about autobiographical events, and each truth teller discussed a different event. We chose the scenario of autobiographical events because truth tellers would have experienced the events and would thus find them meaningful and memorable which in turn enables truth tellers to provide detailed reports (Johnson & Raye, 1981; Sporer & Sharman, 2006). To ensure that the events could be comparable to those of lie tellers, each lie teller was matched to an event a truth teller reported. Therefore, the variability of events may be considered as a strength in our study where matched truth tellers and lie tellers could be compared on the same event, while at the same time generalising the findings across different events.

We cannot establish the ground truth of the reported events (i.e., that truth tellers actually experienced the event in the past two years). Nonetheless, we have no reason to believe that truth tellers lied about the event as they were free to choose the event they wished to talk about. It was in the best interest of truth tellers to appear honest and receive entry into the prize draw. Moreover, when participants were asked to rate their truthfulness in the post-



interview questionnaire, truth tellers indicated that they were overwhelmingly honest (they reported that 96.75% of their report was honest; see Appendix B).

Participants were college students which may have affected the ecological validity of the study as the behaviour of students may differ from that of real suspects. However, a meta-analysis that examined the moderating effects of various factors—such as (mock) suspect motivations, emotions, student vs non-student samples, and interview settings—on differences between truth tellers and lie tellers showed that veracity differences were stable across settings (Hartwig & Bond, 2014). This suggests that findings on deception can be generalised from the lab to real settings.

Moreover, due to ethical reasons, we could not induce stress in the interview to simulate real interview settings. However, we could increase participants' motivation to convince the interviewer they are honest by rewarding them and at the same time telling them that if they are not convincing they will not be entered in a draw and they would need to write a statement about the event. These instructions are commonly used in deception research to simulate participants (e.g., Luke, Hartwig, Shamash, & Granhag, 2016). The manipulation checks (see Appendices) also revealed that participants were highly motivated across interviews.

### *5.5. Conclusions*

The current research showed that the type of detail can distinguish lie tellers from truth tellers. More specifically, complications and the proportion of complications were particularly reliable cues (and showed large effects) for detecting deception and eliciting information, thus replicating previous research. The Model Statement was effective for eliciting common knowledge details over time from both truth tellers and lie tellers when presented in a first interview, and for eliciting new information when presented in a subsequent interview.

**Acknowledgments**

This work was supported by the High-Value Detainee Interrogation Group DJF-18-0017-PR-0006603 awarded to the University of Portsmouth (UK). Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the U.S. Government.

**Declarations of interest**

None.

## References

- Amado, B. G., Arce, R., Fariña, F., & Vilarino, M. (2016). Criteria-Based Content Analysis (CBCA) reality criteria in adults: A meta-analytic review. *International Journal of Clinical and Health Psychology, 16*, 201-210. doi:10.1016/j.ijchp.2016.01.002
- Dando, C., Wilcock, R., & Milne, R. (2009). The cognitive interview: Novice police officers' witness/victim interviewing practices. *Psychology, Crime & Law, 15*, 679-696. doi:10.1080/10683160802203963
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*, 979-995. doi:10.1037/0022-3514.70.5.979
- DePaulo, B. M., Lindsay, J. L., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74-118. doi:10.1037/0033-2909.129.1.74
- Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*. New York City, USA: Columbia University.
- Evans, J. R., Pimentel, P. S., Pena, M. M., & Michael, S. W. (2017). The ability to detect false statements as a function of the type of statement and the language proficiency of the statement provider. *Psychology, Public Policy, and Law, 23*, 290. doi:10.1037/law0000127
- Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., Ivanova, M., Granskaya, J., & Houston, K. (2016). Using the model statement to elicit information and cues to deceit from native speakers, non-native speakers and those talking through an interpreter. *Applied Cognitive Psychology, 30*, 854-862. doi:10.1002/acp.3270
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117-140. doi:10.1177/001872675400700202

- Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*, 2-18. doi:10.1037/a0024338
- Gabbert, F., Hope, L., & Fisher, R. P. (2009). Protecting eyewitness evidence: Examining the efficacy of a Self-Administered Interview tool. *Law and Human Behavior*, *33*, 298-307. doi:10.1007/s10979-008-9146-8.
- Gabbert, F., Hope, L., Fisher, R. P., & Jamieson, K. (2012). Protecting against misleading post-event information with a Self-Administered Interview. *Applied Cognitive Psychology*, *26*, 568-575. doi:10.1002/acp.2828
- Granhag, P.A. & Hartwig, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind-reading. *Psychology, Crime & Law*, *14*, 189-200. doi:10.1080/10683160701645181
- Granhag, P. A., & Strömwall, L. A. (1999). Repeated interrogations—Stretching the deception detection paradigm. *Expert Evidence*, *7*, 163-174. doi:10.1023/A:1008993326434
- Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: Verbal and non-verbal cues to deception. *Applied Cognitive Psychology*, *16*, 243-257. doi:10.1002/acp.784
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*, 23-34. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/>
- Hartwig, M., & Bond, C. F. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology*, *28*, 661-676. doi:10.1002/acp.3052
- Hartwig, M., Granhag, P. A., & Strömwall, L. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime, & Law*, *13*, 213-227. doi:10.1080/10683160600750264

- Harvey, A., Vrij, A., Hope, L., Leal, S., & Mann, S. (2017). A stability bias effect amongst deceivers. *Law and Human Behavior*. doi:10.1037/lhb000025
- Harvey, A., Vrij, A., Leal, S., Hope, L., & Mann, S. (2017). Deception and decay: Verbal lie detection as a function of delay and encoding quality. *Journal of Applied Research in Memory and Cognition*, 6, 306-318. doi: 10.1016/j.jarmac.2017.04.002
- Harvey, A. C., Vrij, A., Leal, S., Lafferty, M., & Nahari, G. (2017). Insurance based lie detection: Enhancing the verifiability approach with a Model Statement component. *Acta Psychologica*, 174, 1-8. doi:10.1016/j.actpsy.2017.01.001
- Heuer, F., & Reisberg, D. (1990). Vivid memories of emotional events: The accuracy of remembered minutiae. *Memory & Cognition*, 18, 496-506. doi:10.3758/BF03198482
- Izovotas, A., Vrij, A., Hope, L., Mann, S., Granhag, P.A., & Strömwall, L. A. (2018). Facilitating memory-based lie detection in immediate and delayed interviewing: The role of mnemonics. *Applied Cognitive Psychology*, 32, 561-574. doi:10.1002/acp.3435
- Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, 7, 2-9. doi:10.7771/1932-6246.1167.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, 88, 67-85. doi:0033-295X/81/8801-0067\$00.75
- Jundi, S., Vrij, A., Hope, L., Mann, S., & Hillman, J. (2013). Establishing evidence through undercover and collective intelligence interviewing. *Psychology, Public Policy, & Law*, 19, 297-306. doi:10.1037/a0033571
- Köhnken, G. (2004). Statement Validity Analysis and the 'detection of the truth'. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 41-63). Cambridge, England: Cambridge University Press.

- Kornell, N., & Björk, R. A. (2009). A stability bias in human memory: overestimating remembering and underestimating learning. *Journal of Experimental Psychology: General*, *138*, 449-468. doi:10.1037/a0017350
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. doi:10.3389/fpsyg.2013.00863
- Leal, S., Vrij, A., Deeb, H., & Jupe, L. (2018). Using the Model Statement to elicit verbal differences between truth tellers and liars: the benefit of examining core and peripheral details. *Journal of Applied Research in Memory and Cognition*, *7*, 610-617. doi:10.1016/j.jarmac.2018.07.001
- Leal, S., Vrij, A., Vernham, Z., Dalton, G., Jupe, L. M., Nahari, G., & Rozmann, N. (2019). Using the model statement to elicit verbal differences between truth tellers and liars amongst Arab interviewees: a partial replication of Leal, Vrij, Deeb and Jupe (2018). *Applied Cognitive Psychology*. Advance online publication. doi:10.1002/acp.3536
- Leal, S., Vrij, A., Warmelink, L., Vernham, Z., & Fisher, R. (2015). You cannot hide your telephone lies: Providing a Model Statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology*, *20*, 129-146. doi:10.1111/lcrp.12017
- Leins, D., Fisher, R. P., & Ross, S. J. (2013). Exploring liars' strategies for creating deceptive reports. *Legal and Criminological Psychology*, *18*, 141-151. doi:10.1111/j.2044-8333.2011.02041.x
- Luke, T. J., Hartwig, M., Shamash, B., & Granhag, P. A. (2016). Countermeasures against the strategic use of evidence technique: Effects on suspects' strategies. *Journal of Investigative Psychology and Offender Profiling*, *13*, 131-147. doi:10.1002/jip.1448

- McDougall, A. J., & Bull, R. (2015). Detecting truth in suspect interviews: The effect of use of evidence (early or gradual) and time delay on Criteria-Based Content Analysis, Reality Monitoring and inconsistency within suspect statements. *Psychology, Crime, & Law*, *21*, 514-530. doi:10.1080/1068316X.2014.994631
- Memon, A., Meissner, C. A., & Fraser, J. (2010). The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law*, *16*, 340. doi:10.1037/a0020518
- Nahari, G. (2018). Reality Monitoring in the forensic context: Digging deeper into the speech of liars. *Journal of Applied Research in Memory and Cognition*, *7*, 432-440. doi:10.1016/j.jarmac.2018.04.003
- Nahari, G., Ashkenazi, T., Fisher, R. P., Granhag, P. A., Hershkowitz, I., Masip, J., ... & Verschuere, B. (2019). 'Language of lies': Urgent issues and prospects in verbal lie detection research. *Legal and Criminological Psychology*, *24*, 1-23. doi:10.1111/lcrp.12148
- Nahari, G., & Vrij, A. (2015) Systematic errors (biases) in applying verbal lie detection tools: richness in detail as a test case. *Crime Psychology Review*, *1*, 98-107. doi:10.1080/23744006.2016.1158509
- Raftery, A. E. (1995). Bayesian model selection in social research. In P. V. Marsden (Ed.), *Sociological methodology* (pp. 111–196). Cambridge, MA: Blackwell.
- Roos af Hjelmsäter, E., Öhman, L., Granhag, P. A., & Vrij, A. (2014). 'Mapping' deception in adolescents: Eliciting cues to deceit through an unanticipated spatial drawing task. *Legal and Criminological Psychology*, *19*, 179-188. doi:10.1111/j.2044-8333.2012.02068.x
- Sakrisvold, M. L., Granhag, P. A., & Mac Giolla, E. (2017). Partners under pressure: Examining the consistency of true and false alibi statements. *Behavioral Sciences & the Law*, *35*, 75-90. doi:10.1002/bsl.2275

- Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working memory model. *Frontiers in Psychology: Hypothesis and Theory*, 7, 420. doi:10.3389/fpsyg.2016.00420.
- Sporer, S. L., & Sharman, S. J. (2006). Should I believe this? Reality monitoring of accounts of self-experienced and invented recent and distant autobiographical events. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20, 837-854. doi:10.1002/acp.1234
- Strömwall, L. A., & Willén, R. M. (2011). Inside criminal minds: Offenders' strategies when lying. *Journal of Investigative Psychology and Offender Profiling*, 8, 271-281. doi:10.1002/jip.148
- Vrij, A. (Ed.) (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester, United Kingdom: John Wiley and Sons.
- Vrij, A. (2016). Baselineing as a lie detection method. *Applied Cognitive Psychology*, 30, 1112-1119. doi:10.1002/acp.3288
- Vrij, A. (2019). Deception and truth detection when analyzing nonverbal and verbal cues. *Applied Cognitive Psychology*, 33, 160-167. doi:10.1002/acp.3457
- Vrij, A., Fisher, R., Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology*, 22, 1-21. doi:10.1111/lcrp.12088
- Vrij, A., Fisher, R. P., Blank, H., Leal, S., & Mann, S. (2016). A cognitive approach to elicit verbal and nonverbal cues to deceit. In J. Van Prooijen, & P. A. M. Van Lange, *Cheating, corruption, and concealment: The roots of dishonesty* (pp. 284-302). Cambridge, United Kingdom: Cambridge University Press.
- Vrij, A., Leal, S., & Fisher, R. P. (2018). Verbal deception and the Model Statement as a lie detection tool. *Frontiers in Psychiatry*, 9, 492. doi:10.3389/fpsyg.2018.00492



- Vrij, A., Leal, S., Fisher, R. P., Mann, S., Deeb, H., Jo, E., Castro Campos, C., & Hamzeh, S. (2019). The efficacy of using countermeasures in a Model Statement interview. *European Journal of Psychology Applied to Legal Context*.
- Vrij, A., Leal, S., Fisher, R. P., Mann, S., Jo, E., Shaboltas, A., ... & Houston, K. (2019). Eliciting information and cues to deceit through sketching in interpreter-based interviews. *Applied Cognitive Psychology*. doi:10.1002/acp.3566
- Vrij, Leal, Jupe, & Harvey (2018). Within-subjects verbal lie detection measures: A comparison between total detail and proportion of complications. *Legal and Criminological Psychology*. doi:10.1111/lcrp.12126
- Vrij, A., Leal, S., Mann, S., Dalton, G. Jo, E., Shaboltas, A., ... & Houston, K. (2017). Using the Model Statement to elicit information and cues to deceit in interpreter-based interviews. *Acta Psychologica*, 177, 44-53. doi:10.1016/j.actpsy.2017.04.011
- Vrij, A., Leal, S., Mann, S., & Fisher, R. P. (2012) Imposing cognitive load to elicit cues to deceit: inducing the reverse order technique naturally. *Psychology, Crime & Law*, 18, 579-594. doi:10.1080/1068316X.2010.515987
- Vrij, A., Leal, S., Mann, S., Fisher, R. P., Dalton, G. Jo, E., Shaboltas, A., Khaleeva, M., Granskaya, J., & Houston, K. (2018). Using unexpected questions to elicit information and cues to deceit in interpreter-based interviews. *Applied Cognitive Psychology*, 32, 94-104. doi:10.1002/acp.3382.
- Vrij, A., Leal, S., Mann, S., Vernham, Z., & Brankaert, F. (2015). Translating theory into practice: Evaluating a cognitive lie detection training workshop. *Journal of Applied Research in Memory and Cognition*, 4, 110-120. doi:10.1016/j.jarmac.2015.02.002
- Vrij, A., Leal, S., Mann, S., Warmelink, L., Granhag, P. A., & Fisher, R. (2010). Drawings as an innovative and successful lie detection tool. *Applied Cognitive Psychology*, 24, 587-594. doi:10.1002/acp.1627

Vrij, A., Mann, S., Jundi, S., Hillman, J. & Hope, L. (2014). Detection of concealment in an information-gathering interview. *Applied Cognitive Psychology, 28*, 860-866.

doi:10.1002/acp.305.

Vrij, A., Meissner, C. A., Fisher, R. P., Kassin, S. M., Morgan III, C. A., & Kleinman, S. M.

(2017). Psychological perspectives on interrogation. *Perspectives on Psychological*

*Science, 12*, 927-955. doi:10.1177/1745691617706515

## THE EFFECTS OF A MODEL STATEMENT

### Appendix A: Pre-Interview Questionnaire Analyses

Due to a technical error, the responses of two participants were missing from Time 1 pre-interview questionnaire and three participants from Time 3 pre-interview questionnaire.

#### Motivation

Three ANOVAs were conducted with motivation to do well during the interview as the dependent variable, and Veracity as factor for Time 1 analysis and Veracity and Model Statement as factors for Time 2 and Time 3 analyses. Results revealed that at Time 1, truth tellers ( $M = 6.25$ ,  $SD = 0.83$ ) were more motivated than lie tellers ( $M = 5.91$ ,  $SD = 1.10$ ),  $F(1, 239) = 7.45$ ,  $p = .007$ ,  $\eta^2 = .03$ . Similarly, truth tellers ( $M = 6.02$ ,  $SD = 0.96$ ) were more motivated than lie tellers ( $M = 5.76$ ,  $SD = 1.09$ ) at Time 2,  $F(1, 235) = 3.96$ ,  $p = .048$ ,  $\eta^2 = .02$ . At Time 3, participants in the MS Week 2 condition were more motivated ( $M = 6.31$ ,  $SD = 0.84$ ) than those in the MS Weeks 1 & 2 ( $M = 5.76$ ,  $SD = 1.15$ ),  $F(3, 232) = 3.53$ ,  $p = .016$ ,  $\eta^2 = .04$ . No other significant effects were found. Thus, with time, participants became more motivated as a function of the Model Statement and not of their Veracity condition. The means also indicate that both truth tellers and lie tellers were highly motivated across interviews.

#### Pre-Interview Preparations

Participants were asked in the pre-interview questionnaires at Times 2 and 3 if they prepared for the interview in the past week. Chi square tests revealed no differences between lie tellers and truth tellers in any of the Model Statement conditions (all  $ps > .094$ ), although at Time 2, lie tellers ( $n = 30$ ) prepared significantly more than truth tellers ( $n = 17$ ) when all Model Statement conditions were combined,  $\chi^2(1) = 3.96$ ,  $p = .047$ , Cramer's  $V = .13$ . The average number of minutes spent preparing prior to Time 2 was 21.96 minutes ( $SD = 18.12$ ), and 24.27 minutes ( $SD = 35.92$ ) prior to Time 3.

#### Preparation Measures

Participants were asked to rate on 7-point scales if their preparation for the interview was sufficient, good, and thorough. The average of these scores was computed for each participant, and three ANOVAs were carried out on this score for Times 1, 2, and 3 data. For Time 1, a one-way ANOVA with Veracity as factor and the preparation measures average score as the dependent variable revealed a significant effect of Veracity,  $F(1, 239) = 12.98$ ,  $p < .001$ ,  $\eta^2 = .05$ . Truth tellers ( $M = 4.86$ ,  $SD = 1.40$ ) rated their preparations significantly higher than lie tellers ( $M = 4.20$ ,  $SD = 1.43$ ). Two Veracity x Model Statement ANOVAs on the preparation measures average score at Times 2 and 3 revealed no significant effects (all  $ps > .071$ ).

### **Preparation Strategies**

Participants were asked whether or not they prepared specific strategies for the interview. At Time 1, a chi square test between Veracity and Strategies revealed that lie tellers ( $n = 64$ ) were significantly more likely to report using a strategy than truth tellers ( $n = 38$ ),  $\chi^2(1) = 10.40$ ,  $p = .001$ , Cramer's  $V = .21$ .

As for the data of Times 2 and 3, two chi square tests between Veracity and Strategies were conducted for each Model Statement condition. The results revealed that at Time 2, lie tellers prepared more strategies than truth tellers in general, but this difference was significant only among those who listened to the Model Statement in Week 1 (20 lie tellers and 7 truth tellers),  $\chi^2(1) = 10.48$ ,  $p = .001$ , Cramer's  $V = .42$ . At Time 3, a significant difference emerged for the MS Week 1 condition (22 lie tellers and 9 truth tellers),  $\chi^2(1) = 10.24$ ,  $p = .001$ , Cramer's  $V = .41$ , for the MS Week 2 condition, (21 lie tellers and 13 truth tellers),  $\chi^2(1) = 4.17$ ,  $p = .041$ , Cramer's  $V = .26$ , and for the MS control condition (19 lie tellers and 9 truth tellers),  $\chi^2(1) = 5.64$ ,  $p = .018$ , Cramer's  $V = .31$ .

Table A.1 shows that the most reported strategy that distinguished truth tellers from lie tellers was reliance on memory, whereas lie tellers were more likely than truth tellers to want

to monitor non-verbal behaviour, rehearse for the interview, use embedded lies, and be simple. At Times 2 and 3, truth tellers—and more so lie tellers—wanted to be consistent with their reports and behaviours at Time 1/ Time 2. Nonetheless, at Times 2 and 3, a few lie tellers mentioned planning to be inconsistent by adding information during the interview to fill in gaps or behaving in a manner that is different than in previous interview(s).

Table A.1

*Frequency of Preparation Strategies Reported by Truth tellers and Lie tellers in the Pre-Interview Questionnaires at Times 1, 2, and 3*

Preparation Strategies	Pre-interview 1		Pre-interview 2		Pre-interview 3	
	T (n=38)	L (n=64)	T (n=38)	L (n=69)	T (n=43)	L (n=80)
Be detailed	19	17	17	13	15	11
Monitor non-verbal behaviour	12	28	11	22	10	25
Rely on memory/Be truthful	10	0	11	0	9	0
Include verifiable details	5	5	4	5	6	2
Rehearse for the interview	5	20	3	19	7	23
Be consistent	3	3	11	23	11	28
Report events chronologically	3	4	2	4	3	1
Be simple	1	4	0	3	2	5
Use embedded lies	0	21	0	20	0	18
Be inconsistent	0	0	0	5	1	3
Be spontaneous	0	0	0	0	0	8

*Note.* The total number in each column may exceed the number of truth tellers/ lie tellers in that column because the same participant may have reported more than one strategy.

Pre-interview 1 = Pre-interview questionnaire at Time 1; Pre-interview 2 = Pre-interview questionnaire at Time 2; Pre-interview 3 = Pre-interview questionnaire at Time 3; T = Truth tellers; L = Lie tellers.

Among participants who did not plan a strategy (Table A.2), truth tellers mostly thought they could rely on their memory or be spontaneous. Lie tellers also wanted to be spontaneous but a few did not know what strategy to prepare. At Times 2 and 3, truth tellers—and more so lie tellers—wanted to be consistent with previous interview(s).

Table A.2

*Frequency of Reasons Provided by Truth tellers and Lie tellers for Not Preparing a Strategy as Reported in the Pre-Interview Questionnaires at Times 1, 2, and 3, and in the Post-Interview Questionnaire*

Non-Preparation reasons	Pre-interview 1		Pre-interview 2		Pre-interview 3		Post-interview	
	T (n=81)	L (n=58)	T (n=82)	L (n=54)	T (n=75)	L (n=42)	T (n=78)	L (n=37)
Rely on memory/Be truthful	70	0	65	0	66	0	71	0
Be spontaneous	18	26	9	22	10	15	18	19
Did not know what to expect	3	9	1	3	2	0	2	0
Did not know of any strategies	0	14	1	11	0	15	0	12
Not good at lying	0	3	0	0	0	0	0	0
Be convincing	0	3	0	0	0	0	0	0
Use an embedded lie	0	1	0	1	0	0	0	0
Want to remain consistent	0	0	5	13	3	10	4	6
Stay calm and not hesitate	0	0	0	1	0	0	0	0

*Note.* The total number in each column may exceed the number of truth tellers/ lie tellers in that column because the same participant may have reported more than one strategy.

Pre-interview 1 = Pre-interview questionnaire at Time 1; Pre-interview 2 = Pre-interview questionnaire at Time 2; Pre-interview 3 = Pre-interview questionnaire at Time 3; T = Truth tellers; L = Lie tellers.

### Appendix B: Post-Interview Questionnaire Analyses

Due to a technical error, the responses of five participants were missing from the post-interview questionnaire.

#### Post-interview Motivation

A two-way ANOVA with Veracity and Model Statement as factors and motivation to perform well in the three interviews as dependent variable revealed no significant effects of Veracity,  $F(1, 230) = 0.71, p = .399, \eta^2 = .003$ , Model Statement,  $F(3, 230) = 1.20, p = .309, \eta^2 = .02$ , or Veracity x Model Statement,  $F(3, 230) = 0.33, p = .805, \eta^2 = .004$ . Both truth tellers ( $M = 5.86, SD = 0.99$ ) and lie tellers ( $M = 5.74, SD = 1.14$ ) were highly motivated.

#### Convincing Strategies

A chi square test between Veracity and Interview Strategies to appear convincing during the three interviews was carried out for each Model Statement condition. Lie tellers ( $N = 84$ ) were significantly more likely than truth tellers ( $N = 39$ ) to report using a strategy,  $\chi^2(1) = 31.02, p < .001$ , Cramer's  $V = .36$ . The strategies were similar to those reported in the pre-interview questionnaires. Table B.1 shows that the most reported strategy that distinguished truth tellers from lie tellers was reliance on memory, whereas lie tellers were more likely than truth tellers to use embedded lie, be (in)consistent, monitor non-verbal behaviour, rehearse for the interview, and be simple.

Table B.1

*Frequency of Convincing Strategies Reported by Truth tellers and Lie tellers in the Post-Interview Questionnaire*

<b>Convincing strategies</b>	<b>Truth tellers (<math>n = 39</math>)</b>	<b>Lie tellers (<math>n = 84</math>)</b>
Be detailed	8	14
Monitor non-verbal behaviour	16	34
Rely on memory/Be truthful	13	0
Include verifiable details	3	6
Rehearse for the interview	5	17
Be consistent	6	17
Report events chronologically	4	3
Be simple	4	9
Use embedded lies	0	21
Be inconsistent	0	7
Be spontaneous	2	7

*Note.* The total number in each column may exceed the number of truth tellers/ lie tellers in that column because the same participant may have reported more than one strategy.

### **Pre-Interview Preparation**

A Veracity x Model Statement MANOVA was carried out on the extent to which participants thought in the past two weeks outside the interview setting about the event or about what they were going to say during the interview. A significant multivariate effect of

Veracity emerged, Pillai's Trace = .06,  $F(2, 229) = 7.04$ ,  $p = .001$ ,  $\eta^2 = .06$ . No significant Model Statement main effect or Veracity x Model Statement interaction effects emerged ( $ps > .457$ ). At the univariate level, a significant effect of Veracity emerged for thinking about what to say  $F(1, 230) = 8.35$ ,  $p = .004$ ,  $\eta^2 = .04$ , but not for thinking about the event,  $F(1, 230) = 0.39$ ,  $p = .535$ ,  $\eta^2 = .002$ . Lie tellers ( $M = 2.79$ ,  $SD = 1.64$ ) were significantly more likely to think about what to say than truth tellers ( $M = 2.24$ ,  $SD = 1.37$ ).

### **Perceived Believability**

We conducted a Veracity x Model Statement MANOVA on the extent to which participants perceived they would be believed by the interviewer, would have to write a statement, or would win a prize. The results revealed a significant multivariate effect of Veracity, Pillai's Trace = .21,  $F(3, 228) = 20.65$ ,  $p < .001$ ,  $\eta^2 = .21$ . No significant Model Statement main effect or Veracity x Model Statement interaction effect emerged ( $ps > .702$ ). At the univariate level, significant effects of Veracity emerged for perceived believability,  $F(1, 230) = 52.30$ ,  $p < .001$ ,  $\eta^2 = .19$ , and likelihood of writing a statement,  $F(1, 230) = 16.86$ ,  $p < .001$ ,  $\eta^2 = .07$ , but not for likelihood of winning a prize,  $F(1, 230) = 0.11$ ,  $p = .742$ ,  $\eta^2 = .00$ . Truth tellers ( $M = 6.31$ ,  $SD = 0.73$ ) were significantly more likely to perceive that they would be believed by the interviewer than lie tellers ( $M = 5.20$ ,  $SD = 1.49$ ). In contrast, lie tellers ( $M = 3.08$ ,  $SD = 1.62$ ) were significantly more likely to think they would have to write a statement than truth tellers ( $M = 2.26$ ,  $SD = 1.48$ ).

### **Model Statement Manipulation**

When participants were asked if they were exposed to a Model Statement in one or more interviews, 24 reported they did not (11 lie tellers and 13 truth tellers) when they actually did. Out of these 24 participants, 14 heard it in Week 1, seven in Week 2, and three in Weeks 1 and 2. Removing these cases did not yield different findings, so they were kept in the final analyses.



A Veracity x Model Statement MANOVA was conducted on the extent to which participants who listened to the Model Statement did so carefully or were thinking about what to say while the Model Statement was played (the MS control condition was not included in this analysis). A significant multivariate effect of Veracity emerged, Pillai's Trace = .05,  $F(2, 150) = 4.22$ ,  $p = .016$ ,  $\eta^2 = .05$ . No significant Model Statement or Veracity x Model Statement effects emerged ( $ps > .236$ ). At the univariate level, Veracity showed an effect on listening carefully to the Model Statement,  $F(1, 151) = 4.36$ ,  $p = .038$ ,  $\eta^2 = .03$ , but not on thinking about what to say,  $F(1, 151) = 2.85$ ,  $p = .093$ ,  $\eta^2 = .02$ . The means showed that truth tellers ( $M = 5.29$ ,  $SD = 1.47$ ) reported to have listened more carefully to the Model Statement than lie tellers ( $M = 4.69$ ,  $SD = 1.87$ ).

A Veracity x Model Statement MANOVA on the extent to which participants reported to have given fewer or more details after listening to the Model Statement (MS control condition not included in the analysis) revealed no significant main effects of Veracity or Model Statement or Veracity x Model Statement interaction effect ( $ps > .077$ ).

### **Interviewer's Characteristics**

We explored whether the interviewer's characteristics may have affected participants' responses. Participants were asked to rate nine items related to the interviewer's characteristics (smooth, bored, satisfied, awkward, engrossed, involved, friendly, active, and positive). An average score was computed for each participant. We examined whether participants differed in their perceptions of these characteristics as a function of Veracity and Model Statement. A Veracity x Model Statement ANOVA was carried out on the average score. The results revealed no significant effects of Veracity,  $F(1, 230) = 0.01$ ,  $p = .932$ ,  $\eta^2 = .00$ , Model Statement,  $F(3, 230) = 0.36$ ,  $p = .781$ ,  $\eta^2 = .01$ , or Veracity x Model Statement,  $F(3, 230) = 0.51$ ,  $p = .674$ ,  $\eta^2 = .01$ .

### **Volunteering Information**

A Veracity x Model Statement ANOVA on the extent to which participants were willing to volunteer information in the three interviews revealed no significant effects of Veracity,  $F(1, 230) = 0.02, p = .895, \eta^2 = .00$ , Model Statement,  $F(3, 230) = 1.54, p = .205, \eta^2 = .02$ , or Veracity x Model Statement,  $F(3, 230) = 0.03, p = .995, \eta^2 = .00$ .

Similarly, a Veracity x Model Statement ANOVA on the extent to which participants decided to limit themselves to reporting only core details of the event to shorten the interviews revealed no significant effects of Veracity,  $F(1, 230) = 0.42, p = .519, \eta^2 = .002$ , Model Statement,  $F(3, 230) = 1.00, p = .394, \eta^2 = .01$ , or Veracity x Model Statement,  $F(3, 230) = 2.02, p = .112, \eta^2 = .03$ .

However, a Veracity x Model Statement MANOVA conducted on the extent to which participants disclosed and concealed information they can remember in the three interviews (percentage scale), revealed a significant multivariate effect of Veracity, Pillai's Trace = .18,  $F(2, 229) = 24.36, p < .001, \eta^2 = .18$ . No significant Model Statement main effect or Veracity x Model Statement interaction effect emerged ( $ps > .219$ ). Truth tellers ( $M = 87.01, SD = 13.66$ ) disclosed significantly more information than lie tellers ( $M = 68.02, SD = 30.35$ ),  $F(1, 230) = 38.44, p < .001, \eta^2 = .14$ . In contrast, lie tellers ( $M = 37.52, SD = 37.20$ ) reported they were significantly more likely to conceal information than truth tellers ( $M = 22.56, SD = 33.89$ ),  $F(1, 230) = 10.57, p = .001, \eta^2 = .04$ .

### **Truthfulness and Deceptiveness**

A Veracity x Model Statement MANOVA on the extent to which participants told the truth or lied in the three interviews (percentage scale) revealed a significant multivariate effect of Veracity, Pillai's Trace = .86,  $F(2, 229) = 711.40, p < .001, \eta^2 = .86$ . No significant Model Statement main effect or Veracity x Model Statement interaction effect emerged ( $ps > .417$ ). At the univariate level, truth tellers ( $M = 96.75, SD = 11.81$ ) were significantly more likely than lie tellers ( $M = 17.02, SD = 23.01$ ) to be truthful,  $F(1, 230) = 1096.58, p$

$< .001$ ,  $\eta^2 = .83$ . Also, lie tellers ( $M = 84.38$ ,  $SD = 21.44$ ) were significantly more likely than truth tellers ( $M = 2.65$ ,  $SD = 12.13$ ) to be deceptive,  $F(1, 230) = 1289.15$ ,  $p < .001$ ,  $\eta^2 = .85$ .

As a manipulation check, participants were asked if their event was genuine or made up. Eight allocated lie tellers reported it was genuine and three allocated truth tellers reported it was deceptive. However, checking pre- and post-interview questionnaire responses of these participants revealed that allocated lie tellers reported using embedded lies and/or being at least 10% deceptive in the interview. Also, truth tellers' open responses revealed that they were honest and provided factual information during the interview. Conducting the analysis without these participants yielded the same results, so they were kept in the analyses.

### **Different Convincing Strategies**

Participants were asked if they used a different strategy than originally planned. A chi square test between Veracity and Different Strategies was conducted for each Model Statement condition. No significant differences emerged across conditions,  $\chi^2(1) = 2.52$ ,  $p = .112$ , Cramer's  $V = .10$ , with a total of 10 lie tellers and four truth tellers reporting using a different strategy. Table B.2 shows that among participants who did not use a different strategy, truth tellers believed they could rely on their memory to appear convincing, whereas lie tellers did not know of a different strategy or wanted to maintain consistency. Truth tellers, and more so lie tellers, did not use a different strategy believing that their previous strategy was effective.

### **Previous Participation in Model Statement Studies**

Participants who previously took part in Model Statement studies did not differ on core, peripheral and common knowledge details, self-handicapping strategies, and complications at Time 3, Pillai's Trace = .05,  $F(10, 474) = 1.10$ ,  $p = .358$ ,  $\eta^2 = .02$ , or on unique details across interviews, Pillai's Trace = .06,  $F(10, 474) = 1.33$ ,  $p = .212$ ,  $\eta^2 = .03$ . Similarly, the proportion scores of complications and core details and the total number of details were not

affected by participation in previous Model Statement studies at Time 3, Pillai's Trace = .02,  $F(6, 478) = 0.93, p = .470, \eta^2 = .01$ , nor were unique details, Pillai's Trace = .02,  $F(6, 478) = 0.88, p = .511, \eta^2 = .01$ .

Table B.2

*Frequency of Reasons Reported by Truth tellers and Lie tellers in the Post-Interview Questionnaire for (Not) Using a Different Convincing Strategy than Originally Planned*

Reasons	Different strategy used		No different strategy used	
	Truth tellers (n=4)	Lie tellers (n=10)	Truth tellers (n=113)	Lie tellers (n=111)
Was detailed	2	0	0	0
Rehearsed for the interview	2	1	0	0
Monitored non-verbal behaviour	1	2	0	0
Acted the same as truth tellers	0	3	0	0
Avoided repetition	0	1	0	0
Elaborated on story	0	1	0	0
Relied on memory/Was truthful	0	0	75	0
Believed my strategy was effective	0	0	27	51
Did not know of a different strategy	0	0	0	29
Wanted to remain consistent	0	0	8	26
Wanted to be spontaneous	0	2	4	5
Did not know what to expect	0	0	2	0